



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07K 14/705, 16/28, C12N 15/12</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/56820</b> <b>(43) International Publication Date:</b> 17 December 1998 (17.12.98)
<b>(21) International Application Number:</b> PCT/US98/12386 <b>(22) International Filing Date:</b> 12 June 1998 (12.06.98)  <b>(30) Priority Data:</b> 60/049,480 12 June 1997 (12.06.97) US  <b>(71) Applicant (for all designated States except US):</b> SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> ELSHOUBAGY, Nabil, A. [US/US]; 1648 Bow Tree Drive, West Chester, PA 19380 (US). LI, Xiaotong [CN/US]; Apartment B41B, 332 Sugartown Road, Devon, PA 19333 (US). BERGSMA, Derk, J. [US/US]; 271 Irish Road, Berwyn, PA 19312 (US). MOONEY, Jeffrey, L. [US/US]; 143 Winged Foot Court, Limerick, PA 19468 (US). GUERRERA, Stephanie, F. [US/US]; Apartment 505, 313 Vine Street, Philadelphia, PA 19106 (US).  <b>(74) Agents:</b> LEWIS, Christopher, R. et al.; Ratner & Prestia, Suite 301, One Westlakes, Berwyn, P.O. Box 980, Valley Forge, PA 19482-0980 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> HM74A RECEPTOR  <b>(57) Abstract</b>  HM74A polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing HM74A polypeptides and polynucleotides in therapy, and diagnostic assays for such.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## HM74A RECEPTOR

### Field of the Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in therapy and in identifying compounds which may be agonists, antagonists  
5 and/or inhibitors which are potentially useful in therapy, and to production of such polypeptides and polynucleotides.

### Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces 'functional genomics', that is, high throughput genome- or gene-based biology. This approach is rapidly  
10 superceding earlier approaches based on 'positional cloning'. A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need  
15 to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, Nature, 1991, 351:353-354). Herein these proteins are referred to as proteins participating in  
20 pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., et al., Proc. Natl Acad. Sci., USA, 1987, 84:46-50; Kobilka, B.K., et al., Science, 1987, 238:650-656; Bunzow, J.R., et al., Nature, 1988, 336:783-787), G-proteins themselves, effector proteins, e.g., phospholipase C, adenylyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M.I., et al.,  
25 Science, 1991, 252:802-8).

For example, in one form of signal transduction, the effect of hormone binding is activation of the enzyme, adenylyl cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide, GTP. GTP also influences hormone binding. A G-protein connects the hormone receptor to adenylyl cyclase. G-protein was shown to exchange GTP for bound GDP when activated by a hormone  
30 receptor. The GTP-carrying form then binds to activated adenylyl cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein

serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

The membrane protein gene superfamily of G-protein coupled receptors has been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane  $\alpha$ -helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors.

G-protein coupled receptors (otherwise known as 7TM receptors) have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. The G-protein family of coupled receptors includes dopamine receptors which bind to neuroleptic drugs used for treating psychotic and neurological disorders. Other examples of members of this family include, but are not limited to, calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins, endothelial differentiation gene-1, rhodopsins, odorant, and cytomegalovirus receptors.

Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 has been implicated in signal transduction.

Phosphorylation and lipidation (palmitoylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxy terminus. For several G-protein coupled receptors, such as the  $\beta$ -adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

For some receptors, the ligand binding sites of G-protein coupled receptors are believed to comprise hydrophilic sockets formed by several G-protein coupled receptor transmembrane domains, said sockets being surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form a polar ligand binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand binding site, such as the TM3 aspartate residue. TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson et al., Endoc. Rev., 1989, 10:317-331).

Different G-protein  $\alpha$ -subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors has been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous sites within a mammalian host. Over the past 15 years, nearly 350 therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market.

### Summary of the Invention

The present invention relates to HM74A, in particular HM74A polypeptides and HM74A polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including the treatment of infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; stroke; ulcers; asthma; allergies; benign prostatic hypertrophy; migraines; vomiting; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation; and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, hereinafter referred to as "the Diseases", amongst others. In a further aspect, the invention relates to methods for identifying agonists and antagonists/inhibitors using the materials provided by the invention, and treating conditions associated with HM74A imbalance with the identified compounds. In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate HM74A activity or levels.

### Description of the Invention

In a first aspect, the present invention relates to HM74A polypeptides. Such peptides include isolated polypeptides comprising an amino acid sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2 over the entire length of SEQ ID NO:2. Such polypeptides include those comprising the amino acid of SEQ ID NO:2.

Further peptides of the present invention include isolated polypeptides in which the amino acid sequence has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:2 over the entire length of SEQ ID NO:2. Such polypeptides include the polypeptide of SEQ ID NO:2.

Further peptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising the sequence contained in SEQ ID NO:1.

Polypeptides of the present invention are believed to be members of the Chemokine family of polypeptides. They are therefore of interest because this invention adds an additional member in the purinergic 7TM receptor family of genes that are involved in a number of biological and disease manifestations. In addition, G protein-coupled receptors, more than any other gene family, are targets of pharmaceutical intervention. These properties are hereinafter referred to as "HM74A activity" or "HM74A polypeptide activity" or "biological activity of HM74A". Also included amongst these activities are antigenic and immunogenic activities of said HM74A polypeptides, in particular the antigenic and immunogenic activities of the polypeptide of SEQ ID NO:2. Preferably, a polypeptide of the present invention exhibits at least one biological activity of HM74A.

The polypeptides of the present invention may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The present invention also includes include variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to HM74A polynucleotides. Such polynucleotides include isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, to the amino acid sequence of SEQ ID NO:2, over the entire length of SEQ ID NO:2. In this regard, polypeptides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and those with at least 99% identity are most

highly preferred. Such polynucleotides include a polynucleotide comprising the nucleotide sequence contained in SEQ ID NO:1 encoding the polypeptide of SEQ ID NO:2.

Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence that has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, to a nucleotide sequence encoding a polypeptide of SEQ ID NO:2, over the entire coding region. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and those with at least 99% identity are most highly preferred.

Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, to SEQ ID NO:1 over the entire length of SEQ ID NO:1. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the polynucleotide of SEQ ID NO:1 as well as the polynucleotide of SEQ ID NO:1.

The invention also provides polynucleotides which are complementary to all the above described polynucleotides.

The nucleotide sequence of SEQ ID NO:1 shows homology with HM74, Nomura,H., Nielsen,B.W. and Matsushima,K.J. Int. Immunol. 5, 1239-1249 (1993). The nucleotide sequence of SEQ ID NO:1 is a cDNA sequence and comprises a polypeptide encoding sequence (nucleotide 61 to 1149) encoding a polypeptide of 363 amino acids, the polypeptide of SEQ ID NO:2. The nucleotide sequence encoding the polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:1 or it may be a sequence other than the one contained in SEQ ID NO:1, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2. The polypeptide of SEQ ID NO:2 is structurally related to other proteins of the Chemokine family, having homology and/or structural similarity with HM74, Nomura,H., Nielsen,B.W. and Matsushima,K.J. Int. Immunol. 5, 1239-1249 (1993).

Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one HM74A activity.

Polynucleotides of the present invention may be obtained, using standard cloning and screening techniques, from a cDNA library derived from mRNA in cells of human placenta DNA, using the

expressed sequence tag (EST) analysis (Adams, M.D., *et al. Science* (1991) 252:1651-1656; Adams, M.D. *et al., Nature*, (1992) 355:632-634; Adams, M.D., *et al., Nature* (1995) 377 Supp:3-174).

Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

5        When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of  
10       the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al., Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

15       Further embodiments of the present invention include polynucleotides encoding polypeptide variants which comprise the amino acid sequence of SEQ ID NO:2 and in which several, for instance from 5 to 10, 1 to 5, 1 to 3, 1 to 2 or 1, amino acid residues are substituted, deleted or added, in any combination.

      Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1, may be used as hybridization probes for cDNA and genomic DNA or as primers for a  
20       nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs from species other than human) that have a high sequence similarity to SEQ ID NO:1. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical, most preferably 95% identical to that of the referent. The probes  
25       or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides.

      A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than human, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the  
30       sequence of SEQ ID NO: 1 or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution



comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes polynucleotides obtainable by screening an appropriate library under stringent hybridization  
5 conditions with a labeled probe having the sequence of SEQ ID NO:1 or a fragment thereof.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is cut short at the 5' end of the cDNA. This is a consequence of reverse transcriptase, an enzyme with inherently low 'processivity' (a measure of the ability of the enzyme to remain attached to the template during the polymerisation reaction),  
10 failing to complete a DNA copy of the mRNA template during 1st strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., PNAS USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech  
15 Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the 'missing' 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers  
20 designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the  
25 design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems which comprise a polynucleotide or  
polynucleotides of the present invention, to host cells which are genetically engineered with such expression  
30 systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Preferred such methods include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *streptococci*, *staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector which is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL* (*supra*). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography.

5 Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the polynucleotide of SEQ ID NO:1  
10 which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine,  
15 saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled HM74A nucleotide sequences. Perfectly matched sequences can be distinguished  
20 from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, e.g., Myers *et al.*, *Science* (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, *Proc Natl Acad Sci*  
25 *USA* (1985) 85: 4397-4401). In another embodiment, an array of oligonucleotides probes comprising HM74A nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, *Science*, Vol 274, pp 610-613 (1996)).

30 The diagnostic assays offer a process for diagnosing or determining a susceptibility to the Diseases through detection of mutation in the HM74A gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression

can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1, or a fragment thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; stroke; ulcers; asthma; allergies; benign prostatic hypertrophy; migraine; vomiting; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation; and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, amongst others.

The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical

Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in  
5 any normal individuals, then the mutation is likely to be the causative agent of the disease.

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

10 Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-  
15 cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, pp. 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention.  
20 Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

Antibodies against polypeptides of the present invention may also be employed to treat the  
25 Diseases, amongst others.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG,  
30 particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood

clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos.  
5 WO94/29458 and WO94/22914.

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response to protect said animal from the Diseases hereinbefore mentioned, amongst others. Yet another aspect of the invention relates to a method of  
10 inducing immunological response in a mammal which comprises, delivering a polypeptide of the present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

A further aspect of the invention relates to an immunological/vaccine formulation  
15 (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a polypeptide of the present invention wherein the composition comprises a polypeptide or polynucleotide of the present invention. The vaccine formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal  
20 injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be  
25 stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

30 Polypeptides of the present invention are responsible for many biological functions, including many disease states, in particular the Diseases hereinbefore mentioned. It is therefore desirable to devise screening methods to identify compounds which stimulate or which inhibit the function of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to

identify those which stimulate or which inhibit the function of the polypeptide. In general, agonists or antagonists may be employed for therapeutic and prophylactic purposes for such Diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. Such agonists, antagonists or inhibitors so-  
5 identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; or may be structural or functional mimetics thereof (see Coligan *et al.*, *Current Protocols in Immunology* 1(2):Chapter 5 (1991)).

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof by means of  
10 a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the  
15 agonist by the presence of the candidate compound is observed. Constitutively active polypeptides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture,  
20 measuring HM74A activity in the mixture, and comparing the HM74A activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and HM74A polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, *J Mol Recognition*, 8:52-58 (1995); and K. Johanson *et al.*, *J Biol Chem*, 270(16):9459-9471 (1995)).

25 One screening technique includes the use of cells which express the receptor of this invention (for example, transfected CHO cells) in a system which measures extracellular pH or intracellular calcium changes caused by receptor activation. In this technique, compounds may be contacted with cells expressing the receptor polypeptide of the present invention. A second messenger response, e.g., signal transduction, pH changes, or changes in calcium level, is then measured to determine whether the potential  
30 compound activates or inhibits the receptor.

Another method involves screening for receptor inhibitors by determining inhibition or stimulation of receptor-mediated cAMP and/or adenylate cyclase accumulation. Such a method involves transfecting a eukaryotic cell with the receptor of this invention to express the receptor on the cell surface. The cell is then

exposed to potential antagonists in the presence of the receptor of this invention. The amount of cAMP accumulation is then measured. If the potential antagonist binds the receptor, and thus inhibits receptor binding, the levels of receptor-mediated cAMP, or adenylate cyclase, activity will be reduced or increased.

Another method for detecting agonists or antagonists for the receptor of the present invention is the  
5 yeast based technology as described in U.S. Patent No. 5,482,835.

The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies  
10 by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

The polypeptide may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to,  
15 ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance,  $^{125}\text{I}$ ), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also  
20 be used to identify agonists and antagonists of the polypeptide which compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as  
25 the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Thus, in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for polypeptides of the present  
30 invention; or compounds which decrease or enhance the production of such polypeptides, which comprises:



- (a) a polypeptide of the present invention;
- (b) a recombinant cell expressing a polypeptide of the present invention;
- (c) a cell membrane expressing a polypeptide of the present invention; or
- (d) antibody to a polypeptide of the present invention;

5 which polypeptide is preferably that of SEQ ID NO:2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

It will be readily appreciated by the skilled artisan that a polypeptide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide, by:

- (a) determining in the first instance the three-dimensional structure of the polypeptide;
- (b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;
- (c) synthesizing candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and
- (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors.

It will be further appreciated that this will normally be an iterative process.

In a further aspect, the present invention provides methods of treating abnormal conditions such as, for instance, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; stroke; ulcers; asthma; allergies; benign prostatic hypertrophy; migraine; vomiting; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation; and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, related to either an excess of, or an under-expression of, HM74A polypeptide activity.

If the activity of the polypeptide is in excess, several approaches are available. One approach comprises administering to a subject in need thereof an inhibitor compound (antagonist) as hereinabove described, optionally in combination with a pharmaceutically acceptable carrier, in an amount effective to inhibit the function of the polypeptide, such as, for example, by blocking the binding of ligands, substrates,

receptors, enzymes, etc., or by inhibiting a second signal, and thereby alleviating the abnormal condition. In another approach, soluble forms of the polypeptides still capable of binding the ligand, substrate, enzymes, receptors, etc. in competition with endogenous polypeptide may be administered. Typical examples of such competitors include fragments of the HM74A polypeptide.

5           In still another approach, expression of the gene encoding endogenous HM74A polypeptide can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered (see, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Alternatively, oligonucleotides which form triple  
10           helices with the gene can be supplied (see, for example, Lee *et al.*, *Nucleic Acids Res* (1979) 6:3073; Cooney *et al.*, *Science* (1988) 241:456; Dervan *et al.*, *Science* (1991) 251:1360). These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

          For treating abnormal conditions related to an under-expression of HM74A and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically  
15           effective amount of a compound which activates a polypeptide of the present invention, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of HM74A by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral  
20           expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For an overview of gene therapy, see Chapter 20, *Gene Therapy and other Molecular Genetic-based Therapeutic*  
25           *Approaches*, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996). Another approach is to administer a therapeutic amount of a polypeptide of the present invention in combination with a suitable pharmaceutical carrier.

          In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide, such as the soluble form of a polypeptide of the present  
30           invention, agonist/antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients

of the aforementioned compositions of the invention. Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, and the like.

The dosage range required depends on the choice of peptide or other compounds of the present invention, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100  $\mu\text{g/kg}$  of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Polynucleotide and polypeptide sequences form a valuable information resource with which to identify further sequences of similar homology. This is most easily facilitated by storing the sequence in a computer readable medium and then using the stored data to search a sequence database using well known searching tools, such as GCC. Accordingly, in a further aspect, the present invention provides for a computer readable medium having stored thereon a polynucleotide comprising the sequence of SEQ ID NO:1 and/or a polypeptide sequence encoded thereby.

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

5 "Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

10 "Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" 15 refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically 20 modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to 25 both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed 30 monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many

types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent  
5 attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing,  
10 phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., Post-translational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF  
15 PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62).

"Variant" refers to a polynucleotide or polypeptide that differs from a reference  
20 polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed  
25 below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the  
30 genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990)). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default  
5 parameters for nucleic acid comparisons.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the  
10 reference sequence of SEQ ID NO:1, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO: 1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference  
15 nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$20 \quad n_n \leq x_n - (x_n \cdot y),$$

wherein  $n_n$  is the number of nucleotide alterations,  $x_n$  is the total number of nucleotides in SEQ ID NO:1,  $y$  is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and  $\cdot$  is the symbol for the multiplication operator, and  
25 wherein any non-integer product of  $x_n$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_n$ . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain  
30 integer number of amino acid alterations as compared to the reference sequence such that the percent

identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein  $n_n$  is the number of amino acid alterations,  $x_n$  is the total number of amino acids in SEQ ID NO:2,  $y$  is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc.,  $\bullet$  is the symbol for the multiplication operator, and wherein any non-integer product of  $x_n$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_n$ .

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO: 2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein  $n_a$  is the number of amino acid alterations,  $x_a$  is the total number of amino acids in SEQ ID NO:2,  $y$  is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and  $\bullet$  is the symbol for the multiplication operator, and



wherein any non-integer product of  $x_a$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_a$ .

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein  $n_a$  is the number of amino acid alterations,  $x_a$  is the total number of amino acids in SEQ ID NO:2,  $y$  is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and  $\bullet$  is the symbol for the multiplication operator, and wherein any non-integer product of  $x_a$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_a$ .

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, e.g., EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

Examples

## Example 1

The HM74 EST was identified from the public database as a potential 7TM receptor. Oligonucleotides (5') were designed at the 5' and the 3' end of the clone. The 5' primer was (CGCCACTTTGCTGGAGCATTCTAGG) (SEQ ID NO:3) and the 3' primer was (AGTTTCCCTAAATCAGATTCTCTGAATC) (SEQ ID NO:4). These oligo's were used to PCR a 1.3 kb 5' fragment using the human placenta cDNA as a template. The PCR fragment was subcloned into pCR2.1 vector and were sequenced. Comparison of the nucleotide sequence of HM74A with the published HM74 revealed 15 nucleotides differences as well as 5 nucleotides insertion at the 3' end of the clone. Further comparison of the amino acid sequence revealed 15 amino acids difference between the HM74A and the published HM74. Furthermore, the 5 nucleotides insertion at the 3' end of the clone resulted in different 3' coding sequence. The HM74A is a 363 amino acids which is 97.7% identical to the published HM74. The cloning procedure was performed twice to confirm the changes in the amino acid sequences. In order to confirm the correct initiation methionine, a cDNA clone containing the entire coding region and the 5' untranslated region was isolated using the human placenta PAPA-cDNA library. Sequence analysis of the clone showed the presence of a stop codon prior to the first initiation methionine.

## Example 1: Mammalian Cell Expression

The receptors of the present invention are expressed in either human embryonic kidney 293 (HEK293) cells or adherent dhfr CHO cells. To maximize receptor expression, typically all 5' and 3' untranslated regions (UTRs) are removed from the receptor cDNA prior to insertion into a pCDN or pCDNA3 vector. The cells are transfected with individual receptor cDNAs by lipofectin and selected in the presence of 400 mg/ml G418. After 3 weeks of selection, individual clones are picked and expanded for further analysis. HEK293 or CHO cells transfected with the vector alone serve as negative controls. To isolate cell lines stably expressing the individual receptors, about 24 clones are typically selected and analyzed by Northern blot analysis. Receptor mRNAs are generally detectable in about 50% of the G418-resistant clones analyzed.

## Example 2 Ligand bank for binding and functional assays.

A bank of over 200 putative receptor ligands has been assembled for screening. The bank comprises: transmitters, hormones and chemokines known to act via a human seven transmembrane (7TM) receptor; naturally occurring compounds which may be putative agonists for a human 7TM receptor, non-mammalian, biologically active peptides for which a mammalian counterpart has not yet been identified; and compounds not found in nature, but which activate 7TM receptors with unknown natural ligands. This

bank is used to initially screen the receptor for known ligands, using both functional (i.e. calcium, cAMP, microphysiometer, oocyte electrophysiology, etc, see below) as well as binding assays.

Example 3: Ligand Binding Assays

Ligand binding assays provide a direct method for ascertaining receptor pharmacology and are adaptable to a high throughput format. The purified ligand for a receptor is radiolabeled to high specific activity (50-2000 Ci/mmol) for binding studies. A determination is then made that the process of radiolabeling does not diminish the activity of the ligand towards its receptor. Assay conditions for buffers, ions, pH and other modulators such as nucleotides are optimized to establish a workable signal to noise ratio for both membrane and whole cell receptor sources. For these assays, specific receptor binding is defined as total associated radioactivity minus the radioactivity measured in the presence of an excess of unlabeled competing ligand. Where possible, more than one competing ligand is used to define residual nonspecific binding.

Example 4: Functional Assay in *Xenopus* Oocytes

Capped RNA transcripts from linearized plasmid templates encoding the receptor cDNAs of the invention are synthesized in vitro with RNA polymerases in accordance with standard procedures. In vitro transcripts are suspended in water at a final concentration of 0.2 mg/ml. Ovarian lobes are removed from adult female toads, Stage V defolliculated oocytes are obtained, and RNA transcripts (10 ng/oocyte) are injected in a 50 nl bolus using a microinjection apparatus. Two electrode voltage clamps are used to measure the currents from individual *Xenopus* oocytes in response to agonist exposure. Recordings are made in Ca<sup>2+</sup> free Barth's medium at room temperature. The *Xenopus* system can be used to screen known ligands and tissue/cell extracts for activating ligands.

Example 5: Microphysiometric Assays

Activation of a wide variety of secondary messenger systems results in extrusion of small amounts of acid from a cell. The acid formed is largely as a result of the increased metabolic activity required to fuel the intracellular signaling process. The pH changes in the media surrounding the cell are very small but are detectable by the CYTOSENSOR microphysiometer (Molecular Devices Ltd., Menlo Park, CA). The CYTOSENSOR is thus capable of detecting the activation of a receptor which is coupled to an energy utilizing intracellular signaling pathway such as the G-protein coupled receptor of the present invention.

Example 6: Extract/Cell Supernatant Screening

A large number of mammalian receptors exist for which there remains, as yet, no cognate activating ligand (agonist). Thus, active ligands for these receptors may not be included within the ligands

banks as identified to date. Accordingly, the 7TM receptor of the invention is also functionally screened (using calcium, cAMP, microphysiometer, oocyte electrophysiology, etc., functional screens) against tissue extracts to identify natural ligands. Extracts that produce positive functional responses can be sequentially subfractionated until an activating ligand is isolated identified.

5    Example 8:      Calcium and cAMP Functional Assays

7TM receptors which are expressed in HEK 293 cells have been shown to be coupled functionally to activation of PLC and calcium mobilization and/or cAMP stimulation or inhibition. Basal calcium levels in the HEK 293 cells in receptor-transfected or vector control cells were observed to be in the normal, 100 nM to 200 nM, range. HEK 293 cells expressing recombinant receptors are loaded with fura 2 and in a  
10   single day > 150 selected ligands or tissue/cell extracts are evaluated for agonist induced calcium mobilization. Similarly, HEK 293 cells expressing recombinant receptors are evaluated for the stimulation or inhibition of cAMP production using standard cAMP quantitation assays. Agonists presenting a calcium transient or cAMP fluctuation are tested in vector control cells to determine if the response is unique to the transfected cells expressing receptor.

## SEQUENCE INFORMATION

## SEQ ID NO:1

CGCCACTTTG CTGGAGCATT CACTAGGCGA GGCCTCCAT CGGACTCACT AGCCGCACTC 60  
ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 120  
5 TTCCGAGATG ACTTCATTGT CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTC 180  
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA 240  
TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTGGCTGACT TTCTACTGAT CATCTGCCTG 300  
CCCTTCCTGA TGGACAACCTA TGTGAGGCGT TGGGACTGGA AGTTTGGGGA CATCCCTTGC 360  
CGGCTGATGC TCTTCATGTT GGCTATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 420  
10 GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC 480  
AATCGGACAG CAGCCATCAT CTCTGCCTT CTGTGGGGCA TCACTATTGG CCTGACAGTC 540  
CACCTCCTGA AGAAGAAGAT GCCGATCCAG AATGGCGGTG CAAATTTGTG CAGCAGCTTC 600  
AGCATCTGCC ATACCTTCCA GTGGCAGGAA GCCATGTTCC TCCTGGAGTT CTTCCTGCCC 660  
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG 720  
15 GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT 780  
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT 840  
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC 900  
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTATTTTTTC CAGCCCATCC 960  
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG 1020  
20 CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1080  
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCTCTTA TCTGGGCCCA 1140  
ACCTCTCCTT AAATAACCAT GCCAAGAAGG GACATTGTCA CCAAGAACCA GCATCTCTGG 1200  
AGAAACAGTT GGGCTGTTGC ATCGAGTAAT GTCAGTGGAC TCGGCCTAAG GTTTCCTGGA 1260  
ACTTCCAGAT TCAGAGAATC TGATTTAGGG AACTAAGCC GAATTCAGC AACTGGCGG 1320  
25 CCGTTACTAG TGGATCCGAG CTCGGTTACC AAGCTTGGCG T 1361

## SEQ ID NO:2

1 MNRHHLQDHF LEIDKKNCCV FRDDFIVKVL PPVLGLEFIF GLLGNGLALW  
51 IFCFHLKSWK SSRIFLFLNLA VADFLLIICL PFLMDNYVRR WDWKFGDIP  
101 RLMLFMLAMN RQGSIIFLTV VAVDRYFRVV HPHHALNKIS NRTAAIISCL  
30 151 LWGITIGLTV HLLKKKMPIQ NGGANLCSSE SICHTFQWHE AMFLEFFLP

- 28 -

201 LGIILFCSAR IIWSLRQRQM DRHAKIKRAI TFIMVVAIVF VICFLPSVVV  
251 RIRIFWLLHT SGTQNCVYR SVDLAFFITL SFTYMNSMLD PVVYFSSPS  
301 FPNFFSTLIN RCLQRKMTGE PDNNRSTSVE LTGDPNKTRG APEALMANS  
351 EPWSPSYLGP TSP

- 29 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

(i) APPLICANT: SMITHKLINE BEECHAM CORPORATION

(ii) TITLE OF THE INVENTION: HM74A RECEPTOR

5 (iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Ratner &amp; Prestia

(B) STREET: P.O. Box 980

(C) CITY: Valley Forge

10 (D) STATE: PA

(E) COUNTRY: USA

(F) ZIP: 19482

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

15 (B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: TO BE ASSIGNED

20 (B) FILING DATE: 12-JUN-1998

(C) CLASSIFICATION: UNKNOWN

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/049,480

(B) FILING DATE: 12-JUN-1997

25 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Prestia, Paul F

(B) REGISTRATION NUMBER: 23,031

(C) REFERENCE/DOCKET NUMBER: GP-70079

(ix) TELECOMMUNICATION INFORMATION:

30 (A) TELEPHONE: 610-407-0700

(B) TELEFAX: 610-407-0701

- 30 -

(C) TELEX: 846169

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1361 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

10	CGCCACTTTG CTGGAGCATT CACTAGGCGA GGCGCTCCAT CGGACTCACT AGCCGCACTC	60
	ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAATAG ACAAGAAGAA CTGCTGTGTG	120
	TTCCGAGATG ACTTCATTGT CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTC	180
	GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA	240
	TCCAGCCGGA TTTTCTGTGT CAACCTGGCA GTGGCTGACT TTCTACTGAT CATCTGCCTG	300
15	CCCTTCCTGA TGGACAATA TGTGAGGCGT TGGGACTGGA AGTTTGGGGA CATCCCTTGC	360
	CGGCTGATGC TCTTCATGTT GGCTATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG	420
	GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC	480
	AATCGGACAG CAGCCATCAT CTCTTGCCCT CTGTGGGGCA TCACTATTGG CCTGACAGTC	540
	CACCTCCTGA AGAAGAAGAT GCCGATCCAG AATGGCGGTG CAAATTGTG CAGCAGCTTC	600
20	AGCATCTGCC ATACCTTCCA GTGGCAGGAA GCCATGTTCC TCCTGGAGTT CTTCTGCCC	660
	CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG	720
	GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT	780
	GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT	840
	TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC	900
25	AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTATTTTTC CAGCCCATCC	960
	TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG	1020
	CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCAACAA AACCAGAGGC	1080
	GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA	1140
	ACCTCTCCTT AAATAACCAT GCCAAGAAGG GACATTGTCA CCAAGAACCA GCATCTCTGG	1200
30	AGAAACAGTT GGGCTGTTGC ATCGAGTAAT GTCAGTGGAC TCGGCCTAAG GTTTCCTGGA	1260
	ACTTCCAGAT TCAGAGAATC TGATTIAGGG AACTAAGCC GAATTCCAGC AACTGGCGG	1320



CCGTTACTAG TGGATCCGAG CTCGGTTACC AAGCTTGGCG T

1361

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 363 amino acids

5 (B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10 Met Asn Arg His His Leu Gln Asp His Phe Leu Glu Ile Asp Lys Lys  
     1                    5                    10                    15  
 Asn Cys Cys Val Phe Arg Asp Asp Phe Ile Val Lys Val Leu Pro Pro  
                     20                    25                    30  
 Val Leu Gly Leu Glu Phe Ile Phe Gly Leu Leu Gly Asn Gly Leu Ala  
 15                    35                    40                    45  
 Leu Trp Ile Phe Cys Phe His Leu Lys Ser Trp Lys Ser Ser Arg Ile  
                     50                    55                    60  
 Phe Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu Ile Ile Cys Leu  
                     65                    70                    75                    80  
 20 Pro Phe Leu Met Asp Asn Tyr Val Arg Arg Trp Asp Trp Lys Phe Gly  
                     85                    90                    95  
 Asp Ile Pro Cys Arg Leu Met Leu Phe Met Leu Ala Met Asn Arg Gln  
                     100                    105                    110  
 Gly Ser Ile Ile Phe Leu Thr Val Val Ala Val Asp Arg Tyr Phe Arg  
 25                    115                    120                    125  
 Val Val His Pro His His Ala Leu Asn Lys Ile Ser Asn Arg Thr Ala  
                     130                    135                    140  
 Ala Ile Ile Ser Cys Leu Leu Trp Gly Ile Thr Ile Gly Leu Thr Val  
                     145                    150                    155                    160  
 30 His Leu Leu Lys Lys Lys Met Pro Ile Gln Asn Gly Gly Ala Asn Leu  
                     165                    170                    175

- 32 -

Cys Ser Ser Phe Ser Ile Cys His Thr Phe Gln Trp His Glu Ala Met  
 180 185 190  
 Phe Leu Leu Glu Phe Phe Leu Pro Leu Gly Ile Ile Leu Phe Cys Ser  
 195 200 205  
 5 Ala Arg Ile Ile Trp Ser Leu Arg Gln Arg Gln Met Asp Arg His Ala  
 210 215 220  
 Lys Ile Lys Arg Ala Ile Thr Phe Ile Met Val Val Ala Ile Val Phe  
 225 230 235 240  
 Val Ile Cys Phe Leu Pro Ser Val Val Val Arg Ile Arg Ile Phe Trp  
 10 245 250 255  
 Leu Leu His Thr Ser Gly Thr Gln Asn Cys Glu Val Tyr Arg Ser Val  
 260 265 270  
 Asp Leu Ala Phe Phe Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met  
 275 280 285  
 15 Leu Asp Pro Val Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Asn Phe  
 290 295 300  
 Phe Ser Thr Leu Ile Asn Arg Cys Leu Gln Arg Lys Met Thr Gly Glu  
 305 310 315 320  
 Pro Asp Asn Asn Arg Ser Thr Ser Val Glu Leu Thr Gly Asp Pro Asn  
 20 325 330 335  
 Lys Thr Arg Gly Ala Pro Glu Ala Leu Met Ala Asn Ser Gly Glu Pro  
 340 345 350  
 Trp Ser Pro Ser Tyr Leu Gly Pro Thr Ser Pro  
 355 360

25 (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- 33 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCCACTTTG CTGGAGCATT CACTAGG

27

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

5

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGTTTCCCTA AATCAGATTC TCTGAATC

28

**What is claimed is:**

- 1 1. An isolated polypeptide selected from the group consisting of:
  - 2 (i) an isolated polypeptide comprising an amino acid sequence selected from the group
  - 3 having at least:
    - 4 (a) 70% identity;
    - 5 (b) 80% identity;
    - 6 (c) 90% identity; or
    - 7 (d) 95% identity
  - 8 to the amino acid sequence of SEQ ID NO:2 over the entire length of SEQ ID NO:2;
  - 9 (ii) an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2 or
  - 10 (iii) an isolated polypeptide which is the amino acid sequence of SEQ ID NO:2.
- 1 2. An isolated polynucleotide selected from the group consisting of:
  - 2 (i) an isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at
  - 3 least
    - 4 (a) 70% identity;
    - 5 (b) 80% identity;
    - 6 (c) 90% identity; or
    - 7 (d) 95% identity;
  - 8 to the amino acid sequence of SEQ ID NO:2, over the entire length of SEQ ID NO:2;
  - 9 (ii) an isolated polynucleotide comprising a nucleotide sequence that has at least:
    - 10 (a) 70% identity
    - 11 (b) 80% identity;
    - 12 (c) 90% identity; or
    - 13 (d) 95% identity;
  - 14 over its entire length to a nucleotide sequence encoding the polypeptide of SEQ ID
  - 15 NO:2;

- 16 (iii) an isolated polynucleotide comprising a nucleotide sequence which has at least:  
17 (a) 70% identity;  
18 (b) 80% identity;  
19 (c) 90% identity; or  
20 (d) 95% identity;  
21 to that of SEQ ID NO: 1 over the entire length of SEQ ID NO:1;  
22 (iv) an isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ  
23 ID NO:2;  
24 (vi) an isolated polynucleotide which is the polynucleotide of SEQ ID NO: 1; or  
25 (vi) an isolated polynucleotide obtainable by screening an appropriate library under stringent  
26 hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or a fragment  
27 thereof;  
28 or a nucleotide sequence complementary to said isolated polynucleotide.
- 1 3. An antibody immunospecific for the polypeptide of claim 1.
- 1 4. A method for the treatment of a subject:  
2 (i) in need of enhanced activity or expression of the polypeptide of claim 1 comprising:  
3 (a) administering to the subject a therapeutically effective amount of an agonist  
4 to said polypeptide; and/or  
5 (b) providing to the subject an isolated polynucleotide comprising a  
6 nucleotide sequence encoding said polypeptide in a form so as to effect production  
7 of said polypeptide activity *in vivo*.; or  
8 (ii) having need to inhibit activity or expression of the polypeptide of claim 1  
9 comprising:  
10 (a) administering to the subject a therapeutically effective amount of an  
11 antagonist to said polypeptide; and/or  
12 (b) administering to the subject a nucleic acid molecule that inhibits the  
13 expression of a nucleotide sequence encoding said polypeptide; and/or

- 14 (c) administering to the subject a therapeutically effective amount of a  
15 polypeptide that competes with said polypeptide for its ligand, substrate, or  
16 receptor.
- 1 5. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression  
2 or activity of the polypeptide of claim 1 in a subject comprising:
- 3 (a) determining the presence or absence of a mutation in the nucleotide sequence  
4 encoding said polypeptide in the genome of said subject; and/or
- 5 (b) analyzing for the presence or amount of said polypeptide expression in a sample  
6 derived from said subject.
- 1 6. A method for screening to identify compounds which stimulate or which inhibit the function of the  
2 polypeptide of claim 1 which comprises a method selected from the group consisting of:
- 3 (a) measuring the binding of a candidate compound to the polypeptide (or to the cells or  
4 membranes bearing the polypeptide) or a fusion protein thereof by means of a label directly  
5 or indirectly associated with the candidate compound;
- 6 (b) measuring the binding of a candidate compound to the polypeptide (or to the cells or  
7 membranes bearing the polypeptide) or a fusion protein thereof in the presence of a labeled  
8 competitor;
- 9 (c) testing whether the candidate compound results in a signal generated by activation or  
10 inhibition of the polypeptide, using detection systems appropriate to the cells or cell  
11 membranes bearing the polypeptide;
- 12 (d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to  
13 form a mixture, measuring activity of the polypeptide in the mixture, and comparing the  
14 activity of the mixture to a standard; or
- 15 (e) detecting the effect of a candidate compound on the production of mRNA encoding said  
16 polypeptide and said polypeptide in cells, using for instance, an ELISA assay.
- 1 7. An agonist or an antagonist of the polypeptide of claim 1.
- 1 8. An expression system comprising a polynucleotide capable of producing a polypeptide of claim 1  
2 when said expression system is present in a compatible host cell.
- 1 9. A process for producing a recombinant host cell comprising transforming or transfecting a cell  
2 with the expression system of claim 8 such that the host cell, under appropriate culture conditions,

3 produces a polypeptide comprising an amino acid sequence having at least 70% identity to the amino  
4 acid sequence of SEQ ID NO:2 over the entire length of SEQ ID NO:2.

1 10. A recombinant host cell produced by the process of claim 9.

1 11. A membrane of a recombinant host cell of claim 10 expressing a polypeptide comprising an  
2 amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO:2 over  
3 the entire length of SEQ ID NO:2.

1 12. A process for producing a polypeptide comprising culturing a host cell of claim 10 under  
2 conditions sufficient for the production of said polypeptide and recovering the polypeptide from the  
3 culture.

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US98/12386

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07K 14/705, 16/28; C12N 15/12

US CL : 435/7.1, 7.2, 69.1, 252.3, 320.1; 530/530, 388.22; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.1, 7.2, 69.1, 252.3, 320.1; 530/530, 388.22; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

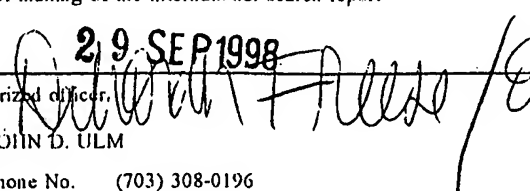
EMBL, GENE BANK, GENESEQ31-2, PIR56, ISSUED PATENTS  
searched SEQ ID NO:1 and 2

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	NOMURA ET AL. Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chemotactic peptide receptors. International Immunology. October 1993. Vol. 5. No. 10. pages 1239-1249, especially HM74 in Figure 2 on page 1241.	1, 2, 8-12 ----- 3-7
X	DEVLIN ET AL. Random Peptide Libraries: A Source of Specific Protein Binding Molecules. SCIENCE. 27 July 1990. Vol. 249. pages 404-406, especially page 404.	7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 28 AUGUST 1998	Date of mailing of the international search report 29 SEP 1998 
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer: JOHN D. ULM Telephone No. (703) 308-0196